Proposed vs. Traditional Metrics in Ranking Generative Models for Cross-Domain Tabular Data

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does the choice of evaluation metric (proposed vs. traditional) influence the ranking of generative models when applied to cross-domain tabular datasets, as assessed through comparative. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: How Evaluation Choices Distort the Outcome of Generative Drug Discovery. Research question: How does the choice of evaluation metric (proposed vs. traditional) influence the ranking of generative models when applied to cross-domain tabular datasets, as assessed through comparative benchmarking on OpenML-CC18 and other standard tabular benchmarks?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.0/10.

3 Results

11 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 7.0/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

References

- http://arxiv.org/abs/2107.12246v2
- http://arxiv.org/abs/2501.05457v2
- http://arxiv.org/abs/2504.20900v1